Ecole du Louvre
ICHIM
Paris 03

www.ichim.org

Les institutions culturelles et le numérique
Cultural institutions and digital technology

LISTEN TO A WEBSITE: ACCESSIBILITY (BEYOND CURRENT STANDARDS) AND A MARKET OPPORTUNITY

Nicoletta Di Blas, Politecnico di Milano, Italy / Paolo Paolin, USI: University of Italian Switzerland & Politecnico di Milano, Italy / Marco Speron, USI: University of Italian Switzerland / Andreas Bienert, Staatliche Museen zu Berlin-Preußischer Kulturbesitz Germany

Abstract

Accessibility to the web, for users with some kind of disabilities, is becoming a major concern for all the institutions, especially those providing resources related to Cultural Heritage. In this paper we focus upon accessibility for those users, with visual disabilities, who utilize a "screen reader", i.e. a SW tool capable of "reading" a web page. In particular this paper illustrates WED, WEb as Dialogue, an innovative approach, developed in a joint effort between HOC-Lab (Politecnico di Milano, Italy), and TEC-Lab (University of Italian Switzerland). The basic idea of WED is that a web experience (i.e. the session of a user with a website) can be assimilated to a dialogue: the user's conversational turns consist in selecting links; the machine's conversational turns consist in providing adequate content (as reply to previous user's requests) and offering new possibilities for the continuation of the dialogue (cues given by links). While a "typical" web page, read via a screen reader, is nearly impossible to understand, the WED approach allows designing websites that offer a more natural "oral interaction" for visually impaired users.
In addition even users with no disabilities may find useful to "listen to" a web site: e.g. while they can't look at a screen or they are using devices with very small screens (as a PDA, for example).
A first concrete result of this technique is the web-site www.munchundberlin.org, prepared (within the project HELP, partially funded by the European Commission), for an exhibition at the Staatliche Museen (Berlin, April 11th - July 13th 2003).

Résumé

L'accessibilité des sites internet, pour les utilisateurs qui ont des déficiences visuelles, est devenue un des plus grands intérêts et une préoccupation pour les institutions, spécialement pour celles qui fournissent des ressources liées au domaine de l'héritage du patrimoine culturel. Cet article se concentre sur l'accessibilité pour les utilisateurs ayant des déficiences visuelles, qui utilisent des synthétiseurs vocaux (i.e logiciel capable de lire les pages des sites web puis de les "exprimer " grâce à une voix de synthèse.). Cet article présente essentiellement WED, « Web as Dialogue », une nouvelle méthodologie innovatrice, développée par le HOC-Lab (Politecnico di Milano, Italie), et le TEC-Lab (Université de la Suisse Italienne). L'idée qui est à la base de WED est qu'une «web experience» (i.e. une session d'usage d'un utilisateur

sur un site web) peut être assimilée à un dialogue: les tours de la conversation de l'utilisateur consistent en une sélection de link ; les tours de la conversation de la machine consistent à fournir des contenus adéquats (en réponse aux requêtes précédentes de l'utilisateur) et à offrir de nouvelles possibilités pour la continuation du dialogue. Alors qu'une page « typique », lue par un synthétiseur vocal, est quasiment impossible à comprendre, la méthodologie WED permet de concevoir des sites web capables d'offrir une « interaction orale » plus naturelle aux utilisateurs ayant des déficiences visuelles.

De plus, même les utilisateurs qui n'ont pas de déficiences visuelles pourraient juger utile d' « écouter » un site web: e.g. alors que ils ne peuvent pas regarder l'écran ou quand ils sont en train d'utiliser un dispositif avec un écran très petit (par exemple un PDA).

Un premier résultat concret de l'application de cette technique est représenté par le site www.munchundberlin.org, financé partiellement par la Commission Européenne, pour une exposition au Staatliche Museen (Berlin, 11 Avril – 13 Juillet 2003)
.

Keywords: web accessibility, web design, dialogue, screen-reader, visual disabilities

Introduction

Many of the activities of our contemporary society are based on information (and possible operations) made available through the World Wide Web. Practical information, images, presentations, sale catalogues, booking and reservations, selling and buying, are exchanged, purchased, activated through the web (or find their vehicle in the web).

The cultural heritage field, meant both in a "traditional" sense (including art, history, archeology…) and in a broad sense (including everyday life, fashion, food…), pays specific attention to the use of this technology: the need for disseminating culture, reaching a wider public, has found in the web one of the best (thanks to its multimedia nature) and cheapest means, much cheaper than print. When the web was born, its core was "hypertext", but soon it turned out to be an eminently "visual" means of communication: its contents are often images and graphics, the page's organization (layout, icons, the contents' and links' position on the page) and the pointing devices (such as the mouse) heavily rely on sight.

There's a non-too small and socially relevant portion of human beings who are affected by some kind of disabilities (physical and/or mental/psychological); this portion can't be left aside by social development: it would be morally unacceptable to keep a large part of society outside of the mainstream opportunity provided by the web (Nevile & McCathieNevile, 2002). In particular, people with visual disabilities (blind or visually impaired people) have difficulties in using the graphic interfaces of modern web applications, while they could perfectly use the character – based interfaces of the older generation.

Accessibility guidelines

The problem of how to make web applications "accessible" to people with visual disabilities has recently gained much interest: W3C consortium, that supplies the "strategic" guidelines for the web, has emanated a standard, based on documents prepared by associations of visually impaired people. The use of a "screen reader", that is, software the "reads" the HTML pages aloud is required; there are many screen readers on the market, with similar characteristics. The WAI ("Web Accessibility Initiative") group of the W3C consortium provides a set of guidelines ("checkpoints") for accessibility; these rules are divided into 14 guidelines, each of them further articulated into sub-points. There are three levels of priority into which the sub-points are organized: priority 1 means basic accessibility (every website should take into consideration guidelines with priority 1); priority 2 means "enhanced" accessibility to content, while priority 3 means taking into consideration specific categories of users with disabilities (http://www.w3.org/TR/WAI-WEBCONTENT). Let's see some example:

1.1 (priority 1): provide a text equivalent for every non-text element (e.g., via "alt", "longdesc", or in element content). This includes: images, graphical representations of text (including symbols), image map regions, animations (e.g., animated GIFs), applets and programmatic objects, ascii art, frames, scripts, images used as list bullets, spacers, graphical buttons, sounds (played with or without user interaction), stand-alone audio files, audio tracks of video, and video.

2.2 (priority 2/3): Ensure that foreground and background color combinations provide sufficient contrast when viewed by someone having color deficits or when viewed on a black and white screen. [Priority 2 for images, Priority 3 for text]

The W3C standard is an important, though still inadequate, step to solve the problem. Even if a site complies with the W3C standard, it is almost unreadable in practice, offering a very cumbersome "conversation" to its user. There are many reasons for this, the most important of which can be listed as follows:

In the web, pages are very complex; someone who can see immediately spots the part of the page s/he's interested in, whilst a blind person has to listen to the whole content before s/he can decide whether there's something that interests her/him.

A web page's organization relies heavily on the graphic; again, someone who can see will immediately access an information "down on the left side", whilst a visually impaired user will have to wait until the voice of the screen reader reaches that point.

The graphic often helps conveying a message: the links' font, color, position, tell the reader about their relevance, their similarities, etc. (a very small link at the bottom of the page, say the "terms and conditions" link, is clearly less relevant than the main links of the page for the average user).

The navigation among pages is usually absolutely inappropriate to engage a sort of effective "dialogue" with its user, who's often compelled to move backward and forward from one page to another, to resume orientation.

Lists of items are practically unusable since again they rely on the graphic; imagine a voice reading a list of 50 different paintings: how could you remember them all? How could you choose one?

The command "back", often used to resume the navigation from a previously-visited node, is very problematic for a user who will have the whole page read again by the screen reader.

For all the above reasons the Politecnico di Milano (HOC Laboratory of the Department of Electronics and Information), in cooperation with TEC-LAB (University of Italian Switzerland), has developed an innovative technology that allows overcoming most of the above listed limitations. The technology is based upon a linguistic approach to the web called WED (WEb as Dialogue; Di Blas & Paolini, 2003a; Di Blas & Paolini, 2003b).

WED interprets the interaction between a human being and a machine in terms of a dialogue; the user's conversation consists of selecting and activating links; the web's conversation consists of providing content (responding to the user's choices) and offering new possibilities of conversation (links).

On the basis of WED new guidelines for accessibility, as well as a new design technique - an evolution of W2000, developed by Politecnico di Milano within UWA EU-funded Project (UWA Consortium, 2001) - have been developed, allowing the implementation of a new generation of websites, very usable and effective for sighted people, optimized for visually impaired people. The first deployment of this technology was the website for the Munch's exhibition that took place in the Staatlichen Museen of Berlin from April 11th to July 13th 2003 (www.munchundberlin.org). Let's see now in details how the WED approach works and what are the main results achieved so far.

The WED approach to web design

It has often been said that it is possible to interpret the interaction between a user and a web application in terms of a dialogue (obviously a very peculiar one!), and that websites might be considered to be "dialogue generators" (Di Blas & Paolini, 2001; Di Blas & Paolini, 2003a). Let us give the name "web-experience" to a user's session with a website: it consists of the sequence of actions (link selection) carried on by the user interleaved with the set of pages that the website displays (reacting to the user's interactions). Starting therefore from the assumption that a website could really improve its effectiveness if well-formed from a dialogic viewpoint, we have translated navigation schemes into dialogue schemes; then we have optimized the dialogue schemes (in general and for the oral channel), and translated the optimized dialogue schemes back into web design.

We focused our attention on information intensive websites, in particular cultural heritage websites, observing many "web-experiences" and trying to make explicit the dialogues between the users and the machine. There follows a short example of transcription of a web experience with the website of the San Francisco Museum of modern Art (www.sfmoma.org); please note that "W" stands for "Web application", whilst "U" stands for "User"; the transcription is made quite freely, adding a bit of "syntactic sugar" in order to make the example clearer; names of paragraphs among square brackets are introduced in order to avoid lengthy repetitions.

W: Welcome to the San Francisco Museum of Modern Art. [T1] You can choose between the following links: visit, info, calendar, education, membership, shop, exhibitions, collections, e-space [end of T1]. As an alternative, you can use the search engine. On evidence we have an exhibition of Marc Chagall (July 26 – November 4, 2003). The news are that you can order advance tickets for Marc Chagall and explore the life and art of Philip Guston in a new multimedia feature. Copyright © 2003 San Francisco Museum of Modern Art. The site is best viewed with Explorer 5.x or better. Do you want to know about the Privacy Policy and the Terms of use?

U: I want to explore the collections.

W: overview of the collections. [T2] You can choose between: recent acquisitions, painting and sculpture, architecture and design, media arts and photography [end of T2]. As an alternative, you can use the search engine. Do you want to search our collections? We are developing an online database of works in our collections that will be available here in summer 2003. In the meantime, thumbnails of many SFMOMA works are available online through AMICO, the Art Museum Image Consortium. The pictures you're seeing are: 1. Brice Marden Cold Mountain 6 (Bridge), 1989-1991; 2. Robert Rauschenberg Collection (formerly Untitled), 1953-1954; 3. T.L. Pflueger Bird's-Eye-View Renderings of the Proposed Airport at China Basin, ca. 1928; 4. Dan Graham

Opposing Mirrors and Video Monitors on Time Delay, 1974; 5. John Gutmann Memory, 1939.
[T1 is repeated]

U: I choose painting and sculpture.

W: [T2 is repeated]. Do you want to use the search engine? Collections: painting and sculpture. The Department of Painting and Sculpture remains true to the Museum's founding vision of presenting substantive and adventurous artistic programming through the acquisition and exhibition of work that reflects the full range of twentieth-century art. You can choose between the following works of art: Jeff Koons, Michael Jackson and Bubbles, 1988; Robert Rauschenberg, Collection(formerly Untitled), 1953-54; Richard Diebenkorn, Ocean Park #54, 1972, Chuck Close, Robert 1997, Anselm Kiefer, Osiris und Isis (Osiris and Isis), 1985-87.
[T1 is repeated]

U: I choose Richard Diebenkorn's Ocean Park.

W: You're exploring the collections: painting and sculpture, and you're seeing the work by Richard Diebenkorn's Ocean Park #54, 1972.
Richard Diebenkorn's work took yet another wide swing in 1967 when he returned to non-objective painting and began a series of monumental abstractions which blended linear geometric structure with fields of luminous color. [more text follows]. Do you want to go back?

U: yes, I want to go back

W: [T2 is repeated]. Do you want to use the search engine? Collections: painting and sculpture. The Department of Painting and Sculpture remains true to the Museum's founding vision of presenting substantive and adventurous artistic programming through the acquisition and exhibition of work that reflects the full range of twentieth-century art. You can choose between the following works of art: Jeff Koons, Michael Jackson and Bubbles, 1988; Robert Rauschenberg, Collection(formerly Untitled), 1953-54; Richard

Diebenkorn, Ocean Park #54, 1972, Chuck Close, Robert 1997, Anselm Kiefer, Osiris und Isis (Osiris and Isis), 1985-87

[T1 is repeated]

Comment: the user wanted to go back to "photography" but s/he couldn't do so directly.

U: I choose photography.

W: …

As the example clearly shows, a dialogue actually takes place: conversational turns alternate, coherence is maintained (that is, the same topics are dealt with by both the interlocutors), a "text" is being created. There are also – of course! – important differences: whilst the user has the power to start/interrupt the conversation whenever s/he wishes, s/he can't contribute with new topics or her/his own ideas, opinions; the so-called semantic dominance (Linell & Luckmann, 1991) is completely on the machine's side. There are some pieces of information (for example, the list of the main links) that are repeated "stubbornly" in every page; this is not a problem if the user is looking at the page (s/he will skip them at a quick glance, if s/he doesn't need them), but if a user is using a screen reader, they make the conversation very cumbersome. Moreover, if a user wants to go back to a "menu" of choices, he has to trace back her/his way, often using the button back of the browser. In a natural conversation, we would rather ask directly our interlocutor for the content we're interested in ("could you tell me about photography?").

We soon understood that all the peculiarities of the "web dialogue" depend basically on two facts:
- one of the interlocutors is a machine
- the communication via web heavily relies on the graphic support, that allows clustering and organizing a lot of content in the same page and conveys part of the message

Therefore we decided to move in two directions: 1) developing a new design technique in order to make the interaction with a web application more "human like" for any kind of

user; 2) trying to optimize the fruition of websites for visually impaired users by making the interaction oral-based rather than graphic-based. This second research line is actually a sprout of the first: once a website is designed having in mind a dialogic approach, what has to be taken into consideration is the way the screen reader reads it: there are very important features that can be implemented in order to achieve this goal, as we shall see below.

Towards optimization

As we said before, accessibility rules allow efficacy but not efficiency of use for visually impaired users, and satisfaction even less. The "screen reader" reads the pages aloud, from top left to bottom right: therefore the graphic support to an almost immediate ("at first glance") understanding of the pages' semantic organization is completely missing; links and content are simply listed one after the other; similar links (sometimes marked by the use of the same color, font) cannot be distinguish from the others, their different relevance (again marked by the use of a smaller font or the position in the page) is not evident.

It was clear that, in order to optimize the web experience for a visually impaired user, further steps were required, and that the comparison with natural dialogues (in particular, as we said before, information intensive dialogues) would be a powerful source of inspiration; in fact, linguistic concepts can be fruitfully used in order to interpret the basic features of the "web dialogue".

In a natural dialogue, we make use of the so-called phoric elements, by which we can refer backward and forward to some elements mentioned in the text, or to the text itself (Bühler 1934; Conte 1989; Cornish 1999). They are called anaphora (reference to something previously mentioned; "can you pass me the salt? I can't see it"), cataphora (reference to something that is about to be mentioned; "I tell you this: I'll never forgive him"), and textual deixis (reference to the text "in that text"; "in the last chapter we talked about the web-dialogue"). We find many counterparts of all these features in the web (of course we have to… enlarge a bit the concept!). Many of the user's moves are anaphoras,

that is, roads back to previously visited pieces of information; when s/he makes use of the button "back" of the browser, s/he making a textual deixis in that s/he wants to go "a step back", that is, s/he's not making a semantic move towards a definite piece of content. Eventually, lists of links can be considered as cataphoras, that is, promises of content still to be told. A user surfing a site actually makes a lot of backward and forward movements in the site, and many of these moves, as usability experts well know, are made because the user is lost and s/he trying to resume the orientation again in the site (and not because s/he's interested in the content of all the pages s/he actually visiting). In a natural conversation, we would never ask our interlocutor to trace back all her/his conversational turns to tell us again something s/he mentioned before. All these moves make the conversation clumsy and cumbersome.

Another feature we paid attention to are long lists of items. The visual support allows a lot of content to be hosted in the page, as the example clearly showed; in a natural conversation, we would never list say all the paintings we know or anything our interlocutor were trying to learn from us, but we would rather try to filter somehow the information.

Moreover, when starting a conversation, we wouldn't immediately enter into the details of every available topic, but we would offer first a reasonable menu of choices ("what shall we talk about?"). In the light of the above observations we started developing new design features, meant to make the web experience more natural and fluid; many of them were explicitly developed having in mind people who can't – for many reasons – look at the screen. These new features were first experimented in the design of the website for an exhibition taking place at the Staatliche Museen of Berlin.

The Munch's exhibition website

The design of the website www.munchundberlin.org represents the first practical result of the WED approach. It complies with – almost – all the accessibility rules of W3C; apart from being accessible, the website presents some features that make it optimized for visually impaired users – and very comfortable for any kind of user (Figure 1).

Figure 1 – www.munchundberlin.or;the homepage

Let's see now in details the basic features of the Munch's site.

Content's sections

The W2000 design methodology - thanks to its concepts and models - has served as a basis for organizing efficiently all the different pieces of information on the page (the main content, the links strictly related to the main content, the links for general services, the landmarks, etc.).

Page schema

The screen reader reads a "page schema" first, that is, a short summary of the basic sections of the page. Therefore the user can immediately choose which section s/he's interested in and, over time, memorize the structure of the pages. The page schema is a purely oral feature: it reflects the page's organization but it is not visualized as text anywhere on the screen.
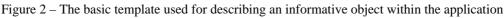
Reading order of the page's content

The screen reader reads the main content of the page first (immediately after reading the "page schema"), even if it is not the first content "graphically" represented on the page.

Page's template

Pages' structure, that is, images, texts and links' positions, remains almost the same in the whole site, enhancing the user's orientation. All the pages are designed according to 2 basic templates (Figure 2).

Figure 2 – The basic template used for describing an informative object within the application

Links' names

The screen reader allows immediate access to the complete list of the page's links. In order to still "mirror" the page's organization ("page schema"), we changed the links' names, specifying the first and last links of each section. In this way, the information about the basic sections of the page is not lost (Figure 3).

Figure 3 – Link's names have further information related to the section they belong

Prints' descriptions

All the descriptions of the prints were prepared trying to keep in mind a user who can't see the prints: the elements of the prints are introduced and then commented.

We further introduced some more features that make the navigation in Munch's site more efficient for any kind of user.

"Visited & missing"

Sometimes a user wonders: "what have I already seen? What's missing?" The "visited & missing" command allows answering to these questions: all the pieces of information already seen or missing are shown in a very efficient synoptic view.

"Just visited"

Sometimes one topic is developed through many pages; therefore, if the user wants to go back to another previously visited topic, s/he normally has to use the "back" of the browser. The "just visited" command aims at offering a shortcut, taking the user directly back to her/his previous "semantic" step.

Up to collection/association centre

This command is similar to the previous one ("just visited"); it allows going back directly to the last "menu" of choices (for example, the last list of prints) visited.

My history in the site

The command "My history in the site" offers a list of the "semantic" steps the user has made so far, allowing a quick re-selection of a previously visited content.

Conclusions and future work

The WED research has already allowed a significant step forward towards the optimization of websites for visually impaired users. We have received very positive feed back from "real life", (visually impaired) users. A German blind user, for example, has written the following email:

"The first impression of the site is very positive. The pages are clearly structured. All the links have detailed titles which allow an informative and nice internet session. With my favourite screenreader JAWS (version 4.51.212) I needed about 1,5 minutes to get a general overview for all further action. This seems to me an acceptable time, considering that this form of documentation of such an exhibition is quite unusual at the moment. Heading with ALT-TAGs - great idea! The engineers did an excellent job by not using the headings only as structural elements, but adding Alt-Tags to them. This creates a very fast and effective structure of orientation which before all will be very convenient for users who surf only occasionally.

To interest sighted people in listening to websites?

The desire of the engineers to interest sighted people in listening to the website is very ambitious, but seems quite hard to achieve as the visitors of the exhibition are living in a visually oriented reality. I tried a few times to convince sighted people to have a try at a computer game without the screen and only presented by sound - but only very few people did so, they were rather an exception. But I like to be proved wrong"

In essence what we have done is to define precise solutions for problems vaguely addressed by the W3C standards; in particular we have tackled those generic recommendations where "good structuring" is suggested, without specific ideas about what to do.

WED instead provides the following innovations:

Definition of an overall design strategy for the whole application
Definition of page schemas optimized for "being read", as parallel with respect to page schemas optimized for "being seen"
Definition of "oral browsing" strategies, allowing a blind user to quickly understand the overall pattern of the application as a whole, of each section and of each page

Future work includes:

Development of systematic/empirical evaluation methodologies, in order to assess more precisely the acceptability of application for blind users.

Definition of guidelines for difficult problems, such as dealing with long lists of items or dealing with text referring to visual experiences.

Improving the effectiveness of navigational patterns for blind users, who can never look at the screen, therefore must rely only upon oral communication.

Development of "semi-oral" navigational patterns, for those users who in a given context (e.g. walking in an archeological park or a museum) would rather listen to the application, but may occasionally look at a (small) screen.

The future work has the overall goal of making the man-machine dialogues (such as those of a user with the web) closer to human-human dialogues (such as those of a user talking with an expert), and their effectiveness.

References

Bühler, K. (1934). Sprachtheorie, Die Darstellungfunktion der Sprache. Jena: Gustav Fischer Verlag

Conte, M.E. (1989, 1977_). Coesione testuale, recenti ricerche italiane. In M.E. Conte (Ed.). La linguistica testuale. Milano: Feltrinelli, 272-295

Cornish, F. (1999). Anaphora, Discourse, and Understanding. Oxford: Clarendon Press

Di Blas, N. (2003). Deixis & Formulations – When the web speaks of himself. SCOMS – Studies in COMmunication Sciences (i.c.s.).

Di Blas, N. & Paolini, P. (2001). La retorica e il web. In P. D'Alessandro (Ed.) Internet e la filosofia. Milano: LED

Di Blas, N. & Paolini, P. (2003a). Do we speak with computers? How linguistics meddle with the web. In E. Rigotti, A. Giacalone-Ramat & A. Rocci (Eds.). Linguistics & new professions. Milano: Franco Angeli (i.c.s.).

Di Blas, N. & Paolini, P. (2003b). "There And Back Again": What Happens To Phoric Elements in a "web Dialogue". Journal of Document Design, 4(3), 194-206

Linell, P. & Luckmann, T. (1991). Asymmetries in Dialogue: Some Conceptual Preliminaries. In I. Markova & K. Foppa (Eds.). Asymmetries in Dialogue. Harvester Wheatsheaf: Hemel Hempstead, 1-20

Nevile, L. & McCathieNevile, C. (2002). The Virtual Ramp to the Equivalent Experience in the Virtual Museum. Accessibility to Museums on the Web. In D. Bearman & J. Trant (Eds.). Museums and the Web 2002. Selected Papers from an International Conference. Pittsburgh: Archives & Museum Informatics, 93-99

UWA Consortium (2001), D7: Hypermedia and Operation Design, Deliverable Project, http://www.uwaproject.org

Online resources:

http://www.w3.org/WAI/: Web Accessibility Initiative
http://www.w3.org/2001/di/: Device Independence Activity
http://www.w3.org/2002/mmi/: Multimodal Interaction Activity
http://www.w3.org/Voice/:"Voice Browser Activity - Voice enabling the Web!